# Percolation of arbitrary words on trees

Ishaan Bhadoo

November 2024

**Abstract**

Consider site percolation on a locally finite, connected graph G satisfying $p_c(G) < 1/2$. The percolation of words problem asks whether all binary sequences can almost surely be embedded in site percolation configurations. Building on a result of Kesten and Benjamini for periodic trees, we demonstrate this for all trees. This is joint work with Ritvik Radhakrishnan.

## 1 Introduction

Consider Bernoulli site percolation on a graph $G$ with $p_c^{\text{site}}(G) < \frac{1}{2}$. Let $\Xi = \{0,1\}^{\mathbb{N}}$ be the set of all infinite binary sequences, referred to as the set of all words. We say that a word $\xi \in \Xi$ is *seen* from a vertex $v \in V$ if there exists a path (infinite self-avoiding path) $v = v_0 \sim v_1 \sim v_2 \sim \cdots$ such that: $\omega_i = \xi_i \quad \forall i \geq 0$.

Define, the set of words seen from a vertex, $S(v) = \{\xi \in \Xi : \xi \text{ is seen from } v\}$. We also define the set of all words seen as $S_\infty = \bigcup_{v \in V} S(v) = \{\xi \in \Xi : \xi \text{ is seen from some } v \in V\}$.

We want study $S_\infty$. We operate on connected, locally finite, (rooted) trees $(T, o)$ satisfying $p_c^{\text{site}}(T) < \frac{1}{2}$ at $p = \frac{1}{2}$. Let us start by stating the main problem we are interested in:

**Conjecture 1** (Classification Conjecture II.(c), [HdNS14])**.** *Let $G$ be an infinite graph, with uniformly bounded degree, and $p_c^{site}$ denote its critical threshold for site percolation. If $p_c^{site} < p < 1 - p_c^{site}$, then all sequences can be embedded, i.e., $\mathbb{P}_p(S_\infty = \Xi) = 1$*

We start by stating Kesten and Benjamini's theorem for periodic trees.

**Theorem 1** (Kesten, Benjamini, [BK95])**.** *Consider a tree $(T, o)$ with period $\nu$, such that $p_c(T) < \frac{1}{2}$. Then $\mathbb{P}_{\frac{1}{2}}(S_\infty = \Xi) = 1$.*

Our main theorem provides a natural extension to this theorem.

**Theorem 2.** *Let $(T, o)$ be a connected locally finite tree such that $p_c^{site}(T) < \frac{1}{2}$. Then $\mathbb{P}(S_\infty = \Xi) = 1$.*

The proof follows the same two steps as the proof by Kesten-Benjamini for periodic trees. The proof is divided into two parts, firstly we show the finiteness of a certain sum and prove that it implies our result, next we show that the sum is indeed finite. The key difference is that we use different events to do the above computation which capture more than just the growth rate of the tree.

Before we dive into the proof, we study site percolation on trees. Define the quantity,

$$\widetilde{br}T = \sup\left\{\lambda : \exists \text{ non-zero flow } \theta \text{ on the vertex set such that } 0 \leq \theta(v) \leq \lambda^{-|v|}\right\}$$

Where flows on vertices is defined in a similar way as for edges, i.e. we have a root vertex $o$ and every vertex other than $o$ satisfies flow conservation. Here $|v|$ is the distance of $v$ from $o$.

By max flow min cut theorem (which holds for countable locally finite graphs), we have that

$$\widetilde{brT} = \sup\left\{\lambda : \inf_{\Pi} \sum_{v\in\Pi} \lambda^{-|v|} > 0\right\}$$

where the infimum is over vertex cutsets $\Pi$, i.e. $\Pi$ is a finite set of vertices such that $o$ is not connected to infinity in $T \setminus \Pi$.

**Lemma 3.** *For site percolation on $T$,*

$$\mathbb{P}_p(o \leftrightarrow \infty) \leq \inf\left\{\sum_{v\in\Pi} p^{|v|} : \Pi \text{ is a (vertex) cutset separating } o \text{ and } \infty\right\}$$

*Proof.* Since, $\{o \leftrightarrow \infty\} \subseteq \bigcup_{v\in\Pi} \{o \leftrightarrow v\}$ for all (vertex) cutsets $\Pi$ ∎

We start by showing $p_c^{\text{site}} \geq \frac{1}{\widetilde{brT}}$. Consider $p > p_c$ and $\lambda = \frac{1}{p}$, then by Lemma 1, for all (vertex) cutsets $\Pi$

$$\sum_{v\in\Pi} \lambda^{-|v|} = \sum_{v\in\Pi} p^{|v|} > 0$$

So, $\widetilde{brT} \geq \frac{1}{p}$ for all $p > p_c$, which implies $p_c^{\text{site}} \geq \frac{1}{\widetilde{brT}}$.

For bond percolation on the tree $T$, it is a famous theorem of Lyons that $p_c^{\text{bond}} = \frac{1}{brT}$. Where $brT$ is the branching number of the tree. By the same argument, a similar statement is true in our case. More specifically, we have:

**Theorem 4.** *Consider site percolation a rooted, locally finite, connected tree $T$. Then*

$$p_c^{site}(T) = \frac{1}{\widetilde{brT}}.$$

We want to show the following theorem,

**Theorem 5.** *Let $T$ be a connected locally finite tree such that $p_c(T) < \frac{1}{2}$. Then $\mathbb{P}(S_\infty = \Xi) = 1$.*

First, we start by setting up some notation: Let $T_n$ denote the set of vertices at distance $n$. For a word $\eta$, let $\eta^{(n)} = (\eta_1, \cdots, \eta_n)$.

For two vertices $v, w \in T$ and an infinite word $\eta \in \Xi$, we denote by $v \xleftrightarrow{\eta^{(n)}} w$ the event that there exists a self-avoiding path $v = v_0 \sim v_1 \sim \cdots \sim v_{n-1} = w$ from $v$ to $w$ along which $\eta^{(n)}$ is seen, i.e., such that $\omega_{v_i} = \eta_i$ for all $i \in \{0, \ldots, n-1\}$. We simply write $\{v \leftrightarrow w\} := \{v \xleftrightarrow{\mathbb{1}^{(n)}} w\}$ for the monochromatic word. For trees, since there is only one path between vertices, for clarity we write (only for trees) $v \overset{\eta}{\leftrightarrow} w$. We also write $x \geq S$ (for a vertex $x$ and set of vertices $S$) if $x$ is a descendant of $S$. We are now ready to show Theorem 8.

*Proof.* Suppose $\widetilde{brT} = 2 + 3\varepsilon$ for some $\varepsilon > 0$. Then by definition of the branching number there exists a non-zero flow $\theta$ such that $0 \leq \theta(v) \leq (2 + 2\varepsilon)^{-n}$. Fix this $\theta$ and let its strength be $||\theta|| = \delta > 0$.

Define random variables,

$$X_n(\eta) = \sum_{v \in T_n} (2+\varepsilon)^n \theta(v) \mathbb{1}(o \overset{\eta}{\longleftrightarrow} v)$$

Also define the events,

$$E_n(\eta) = \{X_n(\eta) \geq \delta\}$$

It is clear that if $E_n(\eta)$ occurs the word $\eta^{(n)}$ must be seen from $o$. For the word $\mathbb{1} = (1, 1, \cdots)$, we set $X_n := X_n(\mathbb{1})$ and $E_n := E_n(\mathbb{1})$.

**Lemma 6.** *If* $\sum_{n \geq 0} 2^n \ \mathbb{P}_{\frac{1}{2}}(E_n \setminus E_{n+1}) < \infty$ *then all words are seen.*

*Proof.* Firstly note that due to symmetry $X_n(\eta) \overset{(d)}{=} X_n(\mathbb{1})$. Thus,

$$\mathbb{P}_{\frac{1}{2}}(E_n(\eta) \setminus E_{n+1}(\eta)) = \mathbb{P}_{\frac{1}{2}}(E_n \setminus E_{n+1}) \ \forall \eta$$

Now suppose that the sum is finite. Consider the events,

$$A_n = \Big\{\text{there exists } \eta^{(n+1)} \text{ such that } E_n(\eta) \setminus E_{n+1}(\eta) \text{ holds}\Big\}$$

We have, $\mathbb{P}(A_n) = \sum_{\eta^{(n+1)}} \mathbb{P}_{\frac{1}{2}}(E_n(\eta) \setminus E_{n+1}(\eta)) = 2^{n+1} \mathbb{P}_{\frac{1}{2}}(E_n \setminus E_{n+1})$, the summability condition this implies that $\sum_{n \geq 0} \mathbb{P}(A_n) < \infty$.

Thus by Borel-Cantelli we get that $\mathbb{P}(A_n \text{ i.o.}) = 0$. So there exists $N = N(\omega)$ such that $A_n$ does not happen a.s. for $n \geq N$ i.e., for all $n \geq N$ and for all $\eta^{(n+1)}$ we have that, $E_n(\eta) \setminus E_{n+1}(\eta)$ does not occur.

Now we show that for all $n \geq 1$, there is a $\eta^{(n)}$ such that $E_n(\eta)$ occurs. Consider

$$\sum_{\eta^{(n)}} X_n(\eta) = \sum_{\eta^{(n)}} \sum_{v \in T_n} (2+\varepsilon)^n \theta(v) \mathbb{1}(o \overset{\eta}{\longleftrightarrow} v) = \sum_{v \in T_n} \sum_{\eta^{(n)}} (2+\varepsilon)^n \theta(v) \mathbb{1}(o \overset{\eta}{\longleftrightarrow} v)$$

But $\sum_{\eta^{(n)}} \mathbb{1}(v \overset{\eta}{\longleftrightarrow} w) = 1$. So,

$$\sum_{\eta^{(n)}} X_n(\eta) = \sum_{v \in T_n} (2+\varepsilon)^n \theta(v) = \delta(2+\varepsilon)^n$$

If $X_n(\eta) < \delta$ for all $\eta^{(n)}$, then the LHS $< \delta 2^n$, which is a contradiction. Thus there must exists a $\eta^{(n)}$ such that $E_n(\eta)$ occurs. Let $n > N$, then by the above observation $E_{n+m}(\eta^{(n)}, \xi^{(m)})$ must also occur for any $\xi$, thus all words are seen from the generation $n$.

$\square$

Now we show that the sum under consideration is indeed finite, in the proof we use McDiarmid's inequality. We start by stating it:

**Theorem 7** (McDiarmid's Inequality)**.** *Let* $X_1, \dots, X_n$ *be independent random variables where* $X_i$ *is* $\mathcal{X}_i$*-valued for all* $i$*, and let* $X = (X_1, \dots, X_n)$*. Assume* $f : \mathcal{X}_1 \times \cdots \times \mathcal{X}_n \to \mathbb{R}$ *is a measurable function such that* $\|D_i f\|_\infty < +\infty$ *for all* $i$*. Then for all* $\beta > 0$*,*

$$\mathbb{P}[f(X) - \mathbb{E}[f(X)] \leq -\beta] \leq \exp\left(-\frac{2\beta^2}{\sum_{i \leq n} \|D_i f\|_\infty^2}\right).$$

*Here, for $x = (x_1, x_2, \cdots x_n)$ the definition of $D_i f(x)$ is:*

$$D_i f(x) := \sup_{y \in \chi_i} f(x_1, x_2, \dots, y, x_{i+1}, \dots, x_n) - \sup_{y' \in \chi_i} f(x_1, x_2, \dots, y', x_{i+1}, \dots, x_n).$$

**Lemma 8.** *There exists $\varepsilon' > 0$ and $C > 0$ such that $\mathbb{P}(E^{\mathsf{c}}_{n+1} | E_n) \leq \exp(-C(1+\varepsilon')^n)$*

It is clear that this lemma implies the summability condition in Lemma 4, so proving Lemma 6 suffices.

*Proof of Lemma 6.* Fix a configuration on $\bigcup_{i \leq n} T_i$ such that $E_n$ occurs. We show the above bound conditioned on this configuration. More formally, let $\tilde{\mathbb{P}}$ be the conditional measure given $E_n, U = \{u_1, \cdots u_k\}$ where $U$ is the set of vertices in $T_n \cap \mathcal{C}$[1] such that $\sum_{1 \leq i \leq k} (2+\varepsilon)^n \theta(u_i) \geq \delta$.

Clearly, $X_{n+1} = \sum_{\substack{|x|=n+1 \\ x \geq U}} (2+\varepsilon)^{n+1} \theta(x) \mathbb{1}(U \leftrightarrow x)$ and thus $\tilde{\mathbb{E}}(X_{n+1}) = \frac{(2+\varepsilon)^{n+1}}{2} \sum_i \theta(u_i)$.

Now, for each descendant of $x$ of $U$ in $T_{n+1}$ consider the random variable $(2+\varepsilon)^{n+1}\theta(x)\mathbb{1}(U \leftrightarrow x) \leq (2+\varepsilon)^{n+1}\theta(x)$. Then by McDiarmid's inequality using $f$ as the sum of these random variables, we get that

$$\mathbb{P}[X_{n+1} - \mathbb{E}[X_{n+1}] \leq -\beta] \leq \exp\left(-\frac{2\beta^2}{\sum_{i \leq n} \|D_i f\|_\infty^2}\right).$$

Now,

$$\sum_i \|D_i f\|_\infty^2 \leq (2(2+\varepsilon)^{n+1})^2 \sum_{\substack{|x|=n+1 \\ x \geq U}} (2+\varepsilon)^{n+1} \theta(x)^2 \leq \frac{4(2+\varepsilon)^{2n+2}}{(2+2\varepsilon)^n} \sum_{i \leq k} \theta(u_i)$$

$$= \frac{4\,(2+\varepsilon)^{n+1}}{(2+2\varepsilon)^n} \cdot 2\,\tilde{\mathbb{E}}(X_{n+1})$$

$$= \frac{8\,(2+\varepsilon)^{n+1}}{(2+2\varepsilon)^n}\,\tilde{\mathbb{E}}(X_{n+1})$$

For $\beta = \frac{\varepsilon}{2+\varepsilon}\tilde{\mathbb{E}}(X_{n+1})$, by McDiarmid,

$$\tilde{\mathbb{P}}\left(X_{n+1} \leq \frac{2}{2+\varepsilon}\tilde{\mathbb{E}}(X_{n+1})\right) \leq \exp\left(-\frac{\left(\frac{\varepsilon}{2+\varepsilon}\right)^2 \tilde{\mathbb{E}}(X_{n+1})\,(2+2\varepsilon)^n}{4\,(2+\varepsilon)^{n+1}}\right).$$

Since $\tilde{\mathbb{E}}(X_{n+1}) \geq \frac{\delta\,(2+\varepsilon)}{2}$,

$$\tilde{\mathbb{P}}\left(X_{n+1} \leq \frac{2}{2+\varepsilon}\tilde{\mathbb{E}}(X_{n+1})\right) \leq \exp\left(-\frac{1}{8}\left(\frac{\varepsilon}{2+\varepsilon}\right)^2 \delta \left(1 + \frac{\varepsilon}{2+\varepsilon}\right)^n\right)$$

$$= \exp\left(-C\left(1+\frac{\varepsilon}{2+\varepsilon}\right)^n\right).$$

[1]cluster of o

Note that,

$$\left\{X_{n+1} \geq \frac{2}{2+\varepsilon}\tilde{\mathbb{E}}(X_{n+1})\right\} \implies E_{n+1}$$

So,

$$\tilde{\mathbb{P}}(E_{n+1} \text{ does not occur}) \leq \tilde{\mathbb{P}}\left(X_{n+1} \leq \frac{2}{2+\varepsilon}\tilde{\mathbb{E}}(X_{n+1})\right) \leq \exp\left(-C\left(1+\varepsilon'\right)^n\right)$$

Where $\varepsilon' = \frac{\varepsilon}{2+\varepsilon}$ and we are done. □

## References


[BK95] Itai Benjamini and Harry Kesten. Percolation of arbitrary words in {0, 1} n. *The Annals of Probability*, pages 1024–1060, 1995.

[HdNS14] M.R. Hilário, B.N.B. de Lima, P. Nolin, and V. Sidoravicius. Embedding binary sequences into bernoulli site percolation on z3. *Stochastic Processes and their Applications*, 124(12):4171–4181, 2014.